

CLAIMS

What is claimed is:

1 1. A method for providing a plurality of client applications access to data in a
2 distributed file system including a plurality of storage elements coupled to a plurality of
3 read servers and a write server, comprising:

4 distinguishing between read requests and write requests from the client
5 applications;

6 transmitting read requests from the client applications to the plurality of read
7 servers;

8 transmitting write requests from the client applications to the write server;

9 reading file data by the read servers and returning the file data to the client
10 applications in response to read requests; and

11 writing data by the write server in response to write requests from the client
12 applications.

1 2. The method of claim 1, in processing a write request at the write server, the
2 write request including a replacement data element to replace a current data element at
3 a current storage location on one of the storage elements, the method further
4 comprising:

5 writing the replacement data element to a new location in the one of the storage
6 elements while leaving the current data element accessible for read requests processed
7 by the read servers;

8 locking a file meta-data element that references the current storage location;
9 updating the file meta-data element to reference the new location having the
10 replacement data element; and

11 unlocking the file meta-data element after updating.

1 3. The method of claim 2, further comprising:

2 allocating a new storage area at the new location prior to writing the
3 replacement data;

4 deallocating an old storage area at the current storage location while the file
5 meta-data element is locked.

1 4. The method of claim 1, further comprising balancing between the read servers a
2 processing load associated with the read requests.

1 5. The method of claim 4, further comprising distributing read requests to the read
2 servers in a round-robin distribution.

1 6. The method of claim 1, further comprising:
2 transmitting read and write requests to the read servers; and
3 transmitting write requests received at the read servers to the write servers.

1 7. A system for providing a plurality of client applications access to data in a
2 distributed file system including a plurality of storage elements, comprising:
3 a plurality of read servers coupled to the client applications and to the
4 distributed file system, each read server configured to read file data from the distributed
5 file system and return the file data to the client applications in response to read requests
6 from the client applications;
7 a write server coupled to the client applications and to the distributed file
8 system, the write server configured to write data to the distributed file system in
9 response to write requests from the client applications;
10 means for distinguishing between read requests and write requests and
11 transmitting read requests from the client applications to the plurality of read servers
12 and transmitting write requests from the client applications to the write server.

1 8. The system of claim 7, wherein a write request includes a replacement data
2 element to replace a current data element at a current storage location on one of the
3 storage elements , the system further comprising:
4 a data consistency control module hosted on the write server, the control
5 module configured and arranged to write the replacement data element to a new
6 location in the one of the storage elements while leaving the current data element
7 accessible to read requests processed by the read servers, lock a file meta-data element
8 that references the current storage location while updating the file meta-data element to
9 reference the new location having the replacement data element, and unlock the file
10 meta-data element after updating.

1 9. The method of claim 7, further comprising means for balancing between the
2 read servers a processing load associated with the read requests.

1 10. The method of claim 9, further comprising means for distributing read requests
2 to the read servers in a round-robin distribution.

1 11. A system for providing a plurality of client applications access to data in a
2 distributed file system including a plurality of storage elements, comprising:

3 a plurality of read servers coupled to the client applications and to the
4 distributed file system, each read server configured to read file data from the distributed
5 file system and return the file data to the client applications in response to read requests
6 from the client applications and configured to transmit write requests to a write server;
7 the write server coupled to the client applications and to the distributed file
8 system, the write server configured to write data to the distributed file system in
9 response to write requests from the client applications;
10 a load balancer coupled to the client applications and to the plurality of read servers, the
11 load balancer configured and arranged to selectively transmit read requests and write
12 requests from the client applications to the plurality of read servers as a function of
13 respective levels of processing loads of the read servers, wherein each of the read
14 servers is further configured and arranged to distinguish between read requests and
15 write requests and transmit write requests to the write server.

1 12. The system of claim 11, wherein a write request includes a replacement data
2 element to replace a current data element at a current storage location on one of the
3 storage elements , the system further comprising:

4 a data consistency control module hosted on the write server, the control
5 module configured and arranged to write the replacement data element to a new
6 location in the one of the storage elements while leaving the current data element
7 accessible to read requests processed by the read servers, lock a file meta-data element
8 that references the current storage location while updating the file meta-data element to
9 reference the new location having the replacement data element, and unlock the file
10 meta-data element after updating.

1 13. The system of claim 12, wherein the load balancer is configured and arranged to
2 distribute the read and write requests to the read servers in round-robin distribution.

1 14. A system for providing a plurality of client applications access to data in a
2 distributed file system including a plurality of storage elements, comprising:
3 a plurality of read servers coupled to the client applications and to the
4 distributed file system, each read server configured to read file data from the distributed
5 file system and return the file data to the client applications in response to read requests
6 from the client applications and configured to transmit write requests to a write server;

7 the write server coupled to the client applications and to the distributed file
8 system, the write server configured to write data to the distributed file system in
9 response to write requests from the client applications;

10 a load balancer coupled to the client applications, to the plurality of read
11 servers, and to the write server, the load balancer configured and arranged to
12 distinguish between read requests and write requests and selectively transmit read
13 requests from the client applications to the plurality of read servers as a function of
14 respective levels of processing loads of the read servers and configured and arranged to
15 transmit write requests to the write server.

1 15. The system of claim 14, wherein a write request includes a replacement data
2 element to replace a current data element at a current storage location on one of the
3 storage elements , the system further comprising:

4 a data consistency control module hosted on the write server, the control
5 module configured and arranged to write the replacement data element to a new
6 location in the one of the storage elements while leaving the current data element
7 accessible to read requests processed by the read servers, lock a file meta-data element
8 that references the current storage location while updating the file meta-data element to
9 reference the new location having the replacement data element, and unlock the file
10 meta-data element after updating.

1 16. The system of claim 14, wherein the load balancer is configured and arranged to
2 distribute the read requests to the read servers in round-robin distribution.

1 17. A system for providing a plurality of client applications access to data in a
2 distributed file system including a plurality of storage elements, comprising:

3 a plurality of read servers coupled to the client applications and to the
4 distributed file system, each read server configured to read file data from the distributed
5 file system and return the file data to the client applications in response to read requests
6 from the client applications and configured to transmit write requests to a write server;
7 the write server coupled to the client applications and to the distributed file
8 system, the write server configured to write data to the distributed file system in
9 response to write requests from the client applications;

10 a plurality of load balancers, each load balancer coupled to a respective one of
11 the client applications, to the plurality of read servers, and to the write server, each load
12 balancer configured and arranged to distinguish between read requests and write
13 requests and selectively transmit read requests from a coupled client application to the
14 plurality of read servers as a function of respective levels of processing loads of the
15 read servers and configured and arranged to transmit write requests to the write server.

1 18. The system of claim 17, wherein a write request includes a replacement data
2 element to replace a current data element at a current storage location on one of the
3 storage elements , the system further comprising:

4 a data consistency control module hosted on the write server, the control
5 module configured and arranged to write the replacement data element to a new
6 location in the one of the storage elements while leaving the current data element
7 accessible to read requests processed by the read servers, lock a file meta-data element
8 that references the current storage location while updating the file meta-data element to
9 reference the new location having the replacement data element, and unlock the file
10 meta-data element after updating.

1 19. The system of claim 17, wherein each of the load balancers is configured and
2 arranged to distribute the read requests to the read servers in round-robin distribution.

1 20. A system for providing a plurality of client applications access to data in a
2 distributed file system including a plurality of storage elements, comprising:

3 a plurality of read servers coupled to the client applications and to the
4 distributed file system, each read server configured to read file data from the distributed

5 file system and return the file data to the client applications in response to read requests
6 from the client applications and configured to transmit write requests to a write server;
7 a write server coupled to the client applications and to the distributed file
8 system, the write server configured to write data to the distributed file system in
9 response to write requests from the client applications;
10 a plurality of load balancers, each load balancer coupled to the plurality of read servers,
11 each load balancer configured and arranged to selectively transmit read requests and
12 write requests from client applications to the plurality of read servers as a function of
13 respective levels of processing loads of the read servers, wherein each of the read
14 servers is further configured and arranged to distinguish between read requests and
15 write requests and transmit write requests to the write server.

1 21. The system of claim 20, wherein a write request includes a replacement data
2 element to replace a current data element at a current storage location on one of the
3 storage elements , the system further comprising:

4 a data consistency control module hosted on the write server, the control
5 module configured and arranged to write the replacement data element to a new
6 location in the one of the storage elements while leaving the current data element
7 accessible to read requests processed by the read servers, lock a file meta-data element
8 that references the current storage location while updating the file meta-data element to
9 reference the new location having the replacement data element, and unlock the file
10 meta-data element after updating.

1 22. The system of claim 20, wherein each of the load balancers is configured and
2 arranged to distribute the read requests and the write requests to the read servers in
3 round-robin distribution.